



[ORIGINAL RESEARCH]

Systematic Study of Structured Diagnostic Procedures in Outpatient Psychiatric Rehabilitation: A Three-year, Three-cohort Study of the Stability of Psychiatric Diagnoses

FUNDING: This research was funded by Skyland Trail.

FINANCIAL DISCLOSURES: Dr. Kotwicki is the Medical Director of Skyland Trail. He reports no other conflicts of interest. Dr. Harvey is a member of the National Advisory board of Skyland Trail and is compensated for this service.

ADDRESS CORRESPONDENCE TO: Raymond Kotwicki, MD; E-mail: rkotwicki@skylandtrail.org

KEY WORDS: Bipolar disorder, psychosis, structured diagnoses, validity

by **RAYMOND KOTWICKI, MD, AND PHILIP D. HARVEY, PhD**

Dr. Kotwicki is with Skyland Trail and Emory University in Atlanta, Georgia; and Dr. Harvey is with the University of Miami Miller School of Medicine, Miami, Florida.

Innov Clin Neurosci. 2013;10(5–6):14–19

ABSTRACT

Background. Psychiatric diagnoses are important for treatment planning. There are a number of current challenges in the area of psychiatric diagnosis with important treatment implications. In this study, we examined the differential usefulness of two semi-structured interviews of differing length compared to clinical diagnoses for generation of diagnoses that did not require modification over the course of treatment.

Methods. We performed a three-year, three-cohort study at an outpatient psychiatric rehabilitation facility, comparing the stability of admission diagnoses when generated by unstructured procedures relying on referring clinician diagnosis, the SCID, and the MINI. We examined changes in diagnoses from admission to discharge (averaging 13 weeks) and, during the second two years, convergence between referring clinician diagnoses and those generated by structured interviews. The same three interviewers

examined all patients in all three phases of the study.

Results. Admission and discharge diagnoses were available for 313 cases. Diagnoses generated with the unstructured procedure were changed by discharge 74 percent of the time, compared to four percent for SCID diagnoses and 11 percent for MINI diagnoses. Referring clinician diagnoses were disconfirmed in Years 2 and 3 in 56 percent of SCID cases and 44 percent of MINI cases. The distinctions between unipolar and bipolar disorders were particular points of disagreement, with similar rates of under and over-diagnosis of bipolar disorder. The rate of confirmation of referring clinician diagnoses of schizoaffective disorder was 10 percent with the SCID and 11 percent with the MINI.

Discussion. In this setting, there appears to be a reasonable trade-off between brevity and accuracy through the use of the MINI compared to the SCID, with substantial improvements in stability of diagnoses compared to clinician

diagnoses. Clinical diagnoses were minimally overlapping with the results of structured diagnoses, suggesting that structured assessment, particularly early in the illness or in short term treatment settings, may improve treatment planning.

INTRODUCTION

The reliability of psychiatric diagnoses has improved markedly since the introduction of structured psychiatric interviews.¹ These interviews were first developed in the late 1960s² and were fine tuned³ up through the time of the introduction of the the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition* (DSM-III)⁴ in 1980. At the same time, the use of these structured interviews is still not common in everyday clinical practice, with most use in research settings. It is not clear how much the application of such interviews would impact the reliability and validity of diagnoses in clinical practice settings, but it seems likely that there are certain circumstances where the increase in validity would be quite substantial. The importance of collection of valid assessment data through structured assessment procedures is compounded by the problems in self-report seen in multiple psychiatric conditions;⁵⁻⁷ questionnaire or checklist methods that do not contain interaction and observation with an interviewer are clearly subject to these concerns.

While we have recently shown in a literature review⁸ that established schizophrenia can be diagnosed by clinicians with high degrees of concordance with the results of structured psychiatric interviews, there are still multiple diagnostic challenges. Patients with multiple, early-course conditions, even schizophrenia, often have diagnoses that change even when initially generated with structured procedures.^{9,10} Psychiatric interviews vary in their focus (Axis-I vs. Axis-II), in their length, and in their assessment of the patient alone

versus symptoms in their relatives. Structured interviews can require substantial time commitments and can require considerable training in order to be accurately employed.

Secular trends and patient expectations may also impact presumed diagnoses when new patients present for treatment in community mental health settings. Some of this variation may be due to exposure of potential patients to media or internet information, which may shape their opinions of their diagnoses. Bipolar disorder, for instance, has seen a marked increase in terms of its diagnosed prevalence in the last 20 years, after 40 years of stability in diagnostic prevalence,¹¹ with this increase corresponding with multiple, newly indicated treatments and associated advertising. In addition, an increased appreciation of the fact that bipolar disorders can be marked by brief episodes of hypomania rather than full manic episodes has increased the challenge in discrimination between bipolar and unipolar mood disorders. We know that distinguishing unipolar depression and bipolar illness has socioeconomic and functional implications.¹² Correspondingly, contemporary diagnostic trends may also incorrectly shape referring diagnoses when patients initially present for treatment. For instance, in previous years the concept of schizophrenia was expanded to include a variety of conditions outside the current boundaries, such as “pseudo-neurotic schizophrenia;”¹³ there is a controversy about whether current concepts of mood spectrum conditions are overly broad as well.

There are several benefits of systematic collection of diagnostic data in everyday practice. There are suggestions that certain conditions, such as bipolar disorder, are both over-diagnosed^{14,15} and frequently missed^{16,17} in clinical settings. The most frequent suggestion to remedy this situation is a structured psychiatric interview. In fact, in the Pogge et al¹⁴ and Zimmerman et al¹⁵ studies, using a structured interview

revealed over-diagnosis of bipolar disorder in adolescent and adults found to have major depression. Presumptive diagnoses of posttraumatic stress disorder (PTSD) are often generated on the basis of trauma exposure, without a systematic assessment of the other required symptoms.¹⁸ Schizoaffective disorder is commonly diagnosed in clinical practice,¹⁹ but the diagnosis has been argued to lack reliability²⁰ and intrinsic clinical validity.²¹

Managed care companies are often interested in matching treatments to diagnoses and may refuse to reimburse for treatments that are not approved for specific indications, suggesting that in order to offer suitable treatments to patients accurate diagnosis is important. This is particularly relevant to time-limited treatment. As interventions such as day treatment or other rehabilitation therapies may be approved by insurance payers for delivery only for finite periods, inaccurate targeting of treatment interventions early on could lead to therapeutic interventions being applied for relatively abbreviated and potentially inefficacious periods. Thus, early identification of the eventual diagnosis can lead to enhanced ability to deliver appropriate treatments for a larger proportion of the time allowed. In this context, stability of diagnoses over time reflects an important component of the validity of these diagnoses while it is admittedly not the only important aspect.

This paper presents the results of a systematic study of the usefulness of structured psychiatric interviews. In a three-year, three-cohort, consecutive-admission study, we examined psychiatric diagnoses that were generated through unstructured clinical interviews and reliance on referral source diagnoses (Year 1), and two different psychiatric interviews that varied in their length of administration (Years 2 and 3). This study was performed at an outpatient psychiatric rehabilitation center that largely focuses on early

TABLE 1. Diagnoses at admission over a three-year, three-cohort prospective study

N	DIAGNOSTIC METHOD (YEAR)					
	CLINICAL (ONE)		SCID (TWO)		MINI (THREE)	
	110		101		110	
	M	SD	M	SD	M	SD
Age	36.4	12.6	34.1	13.4	31.9	11.8
Length of stay (days)	90.6	66.6	86.7	75.2	81.8	51.2
Gender (% male)	62		59		50	
Axis-I admission diagnosis						
Bipolar	46		38		50	
Major depression	30		38		40	
Schizoaffective	11		1		1	
Schizophrenia	6		16		7	
Anxiety (includes PTSD)	6		4		10	
Other	1		4		2	

course patients (mean age=24) and included three years of consecutive admissions from similar referral sources, where the assessment procedure was systematically changed at one-year periods with the same admission staff in place across the three years. We used the Structured Clinical Interview for the DSM (SCID)²² for the second year of the study and the MINI International Neuropsychiatric Interview (MINI)²³ for the third. Stability of diagnoses was indexed through the number of changes in diagnosis suggested by the clinical staff during the course of the patient's treatment based on real-time observations and the results of the therapeutic process. For cases in Years 2 and 3, we also compared the referral source diagnosis for the patients to the diagnosis generated with a structured psychiatric interview. Our hypothesis was that both of the structured interviews would be superior for generating stable diagnoses to both clinical judgments and referral diagnoses based on unstructured clinical observation. We were particularly interested in whether the considerably more abbreviated MINI would yield the same diagnostic

stability, compared to the lengthier SCID, in these patients.

METHODS

Participants. Research participants consisted of three years of consecutive admissions to a private, nonprofit, psychiatric rehabilitation facility. All admissions were examined; cases who were screened for admission but who did not receive services were not analyzed. All data were archived in a database and examined anonymously. Patients signed a general consent form for their data to be examined anonymously and the Emory University Internal Review Board approved this study with expedited review and did not require signed informed consent for the analyses performed in this study. Patients with a primary diagnosis of a substance use disorder or personality disorders were excluded from admission due to regulatory issues during this time period. Dual-diagnoses patients as well as patients who had concomitant (but not primary) personality disorders were included in analyses.

The same three experienced, master-level, admission staff members participated over all three

years. At the beginning of the study, these staff members had a minimum of three years of experience and an average of 5. Cases were distributed sequentially across the three raters after referral to the treatment facility. These staff members were not involved in the treatment of the patients and did not have input into any subsequent treatment decisions. Further, the clinical staff members treating the patients were not informed of the plans to evaluate diagnostic stability as an outcome measure in the study. The reporting of the diagnoses consisted of the axis I and axis II diagnostic impressions which were entered into the electronic medical records. For this study, we focused on axis I diagnoses, as they were primary. Demographic data, including admission diagnoses, are presented in Table 1. As can be seen in the table, the ages of the cases declined slightly each year and there was a slight shift in the diagnostic distribution.

Procedure. The same three admission staff members participated in all three years of this study, which started October 1, 2008. In year one, all referrals for admission to the treatment center received a clinical diagnosis based on an interview at admission and information provided by the referral source. Throughout that year, the presumed "working diagnosis" was the referral diagnosis accompanied by an unstructured interview that occurred within 48 hours of the patients' admission. In a pre-planned study, the three staff members were trained by an experienced psychiatric diagnostician. During Year 2, these same staff members, after training, interviewed all candidates for admission with the SCID. Interview training consisted of observed interviews, joint ratings, and consensus discussion of a series of cases not included in these analyses. After one year of use of the SCID, a third year of admissions were all interviewed and diagnosed with the MINI interview using identical training procedures.

There are several other important features of this design. In order for the procedure to simulate the reality of clinical practice, we did not perform extensive assessment of inter-rater reliability after the initiation of the project. Instead, our goal was to determine if using a semi-structured interview and a trained rater would generate stable diagnoses. To examine this question, we compared the rate of clinician change of rater-generated admission diagnoses across the three raters. Thus, the outcome was diagnostic stability across raters within rating method and not agreement on a specific diagnosis such as bipolar disorder or schizophrenia. Patients were treated at this facility on average over three months, with regular therapist and psychiatric consultations and round-the-clock clinical observation.

Primary axis-I diagnoses were examined during the entire period of treatment for each case during the three year period. Changes in the original admission diagnosis prior to discharge from treatment were recorded as the primary outcome measure. As a secondary outcome in Years 2 and 3, the original clinician diagnosis was compared to the admission diagnosis assigned following the structured diagnostic interview. In the calculation of “change in diagnosis,” we used the following rules: 1) We generated global categories in order to avoid characterizing minor changes in diagnoses as discrepant. For instance, we considered a diagnosis of bipolar II and bipolar I disorder to be consistent, although a change in diagnosis from bipolar depression to major depression was considered a change. 2) We did not consider schizophrenia subtypes as part of the diagnostic agreement, but considered schizophrenia to be different from schizoaffective disorder. 3) Changes in clinical state codes (i.e., severe to remission) during the course of treatment within the same diagnosis were not considered as a difference in diagnosis.

TABLE 2. Changes in referral source diagnoses across during Years 2 and 3 of the study after application of structured diagnostic interviews

REFERRAL DIAGNOSIS	YEAR 2, % CONFIRMED	YEAR 3, % CONFIRMED
Bipolar	40	50
Major depression	50	65
Schizoaffective	10	11
Schizophrenia	50	57
Anxiety	75	50
Overall confirmation rate	44	58

RESULTS

As can be seen in Table 1, the overall pattern of admission diagnoses changed slightly over the study period. Diagnoses of major depressive disorder became more common and diagnoses of psychosis became less so. When the primary outcome, change in diagnosis from admission to discharge, was examined there were clear differences across the methods. In the year prior to the implementation of the SCID, 74 percent of admission diagnoses based on referral diagnosis and unstructured interview were changed over the course of the treatment period, with the most common changes being that an admission diagnosis of major depression was changed to a diagnosis of bipolar disorder or vice versa. In marked contrast, the rates of change of diagnoses generated by the SCID during similarly lengthy stays was four percent and the rates of changes in MINI diagnoses was 11 percent. Chi-square tests were used to compare the differences in rate of change in diagnoses across procedures. The difference in rates of change between the clinical diagnostic assessment procedure and the SCID was significant, $\chi^2(1)=19.09, p<0.001$, as was the difference between clinical diagnoses and the MINI, $\chi^2(1)=7.50, p<0.005$. However, the difference in rates of

changes in diagnoses between SCID and MINI procedures was not significant, $\chi^2(1)=2.01, p=0.16$.

In order to determine whether there were differences across the three raters in the extent to which their diagnoses were changed over the course of treatment for the patients, we performed a 3 (rater) x 2 (changed or not changed) Chi-square test for each of the three years. All three years suggested no differences across the three raters in the extent to which their diagnoses were changed by the clinicians (all $\chi^2(2 df)<1.47$, all $p>0.48$).

In the analyses of data from Years 2 and 3, we compared clinical diagnosis provided by the referring source to the diagnoses generated with the structured procedures. These data are presented in Table 2. See Figure 1 for a graphic depiction of these results. There were substantial discrepancies between these diagnoses. Of the cases interviewed with the SCID, 56 percent of the cases were assigned a diagnosis that was different from that provided by the referral source and for the MINI the number of cases whose diagnosis was different was 42 percent. Diagnostic confirmation rates for bipolar were 40 percent and 50 percent for the two years, and confirmation of major depressive disorder were somewhat higher. Most diagnostic discrepancies were

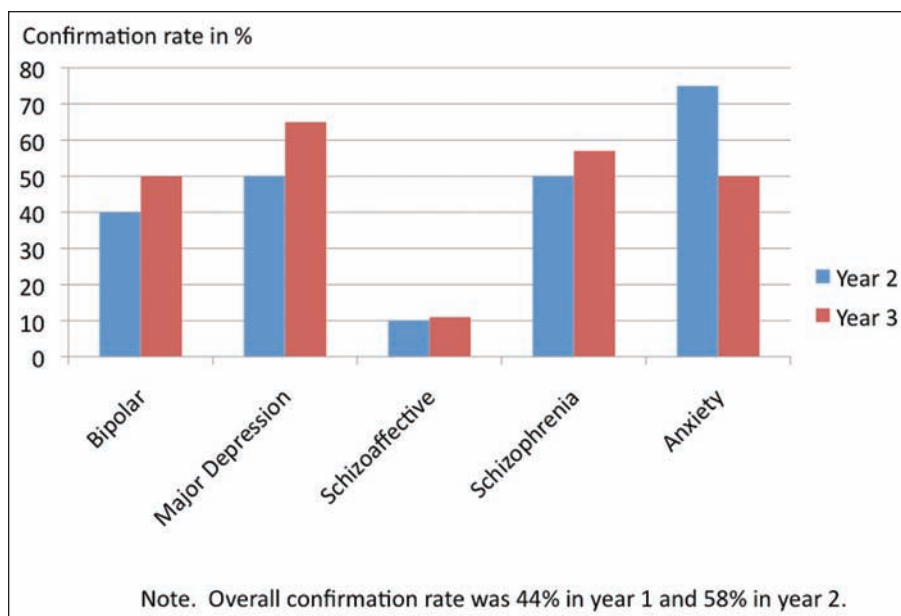


FIGURE 1. Confirmation of referral source diagnoses across during Years 2 and 3 of the study after application of structured diagnostic interviews

interview-based diagnoses of major depression in cases referred as bipolar and bipolar depression in cases referred as major depression. Diagnostic confirmation of schizoaffective disorder was also very low at 10 percent and 11 percent across the two years.

DISCUSSION

Given the importance of matching appropriate diagnosis with evidence-based pharmacologic, psychotherapeutic, and psychosocial interventions, identifying an accurate working diagnosis quickly and efficiently in community mental health settings is essential. Relying on referring diagnoses and self-report of previous diagnoses may yield a diagnosis that requires modification, even in the context of an initial unstructured psychiatric diagnostic assessment by experts. Implementing structured interviews in such settings may be prudent, as rates of diagnostic changes were significantly reduced in our study using both the SCID and MINI. Length of administration of the MINI is shorter than administering the SCID (20 minutes compared to over 90 minutes), and diagnostic stability was not notably different between

these two structured interviews in our cohorts.

Themes of modifications in referral diagnoses in our study mirrored data from other community diagnostic studies. In our Year-2 and Year-3 cohorts, patients who presented initially with diagnoses of bipolar illness, unipolar depression, and schizoaffective disorder had significant rates of re-diagnosis of their primary mental illness using a structured diagnostic tool. Bipolar illness prior to admission seemed to be both over-diagnosed and under-recognized within patients referred to this treatment facility. While the origin of the pattern of such diagnostic changes is unclear, there was both a bias toward over-diagnosis of bipolar disorder in the absence of symptomatic evidence in patients eventually diagnosed with major depression and tendencies to miss euphoric bipolar symptoms. Unsystematic assessment may produce both types of diagnostic errors: inadequate knowledge of the signs of a manic or hypomanic episode may lead to a false positive bipolar diagnosis and failure to assess for manic episodes may lead to false negatives. Schizoaffective disorder seems to be more often

found in the diagnostic opinions of clinicians than in the results of structured assessments.

There are some limitations to these data and these analyses. Stability is not the only element of validity; we did not examine treatment response, biomarkers, or course of illness as validity indicators. We could not quantify the reasons why clinicians changed admission diagnoses, and there may be several reasons for these changes. It also is possible that clinicians were less likely to change diagnoses generated by structured interviews, but the clinicians were not aware we planned to examine changes in diagnosis as an outcome variable in a research study. Discrepancies between referring clinician diagnoses and the results of structured assessments cannot be attributed to bias on the part of clinicians toward not modifying a diagnosis. The facility at which these analyses occurred is a private, non-profit treatment program that does not bill Medicaid or Medicare. A selection bias in referred patients may limit generalizability of these outcomes to other treatment settings in which patients from more varied economic groups are assessed and the applicability to inpatient settings cannot be determined. The very short stays typical in current inpatient treatment make the use of structured diagnostic assessments less useful. A relatively younger age of the patients in this study may limit generalizability of the findings to older patients who have had longer experiences with serious, persistent mental illnesses. Neither the SCID nor the MINI was designed for the purpose of diagnosing axis-II pathology and as a result, these diagnoses could not be systematically assessed in this study. The required investment in training clinicians and administering the standardized assessments may similarly limit the practicality of assessment of these outcomes in busy community mental health treatment centers, where resources tend to be limited.

Although the study participants were on average quite young, patients were typically not recovering from their first episode of illness, during which time actual changes in symptomatology and presentation might account for diagnostic uncertainty. As treatment options for mental illnesses continue to improve, diagnostic stability and reliability become even more important in community mental health settings. Pharmacologic, psychotherapeutic, and social interventions used to treat patients with bipolar illness are significantly different than similar classes of interventions for patients with personality disorders or even unipolar depression. This study suggests that the up-front investment of effort and time to use a structured diagnostic assessment at the time of admission to residential, partial hospitalization, and intensive outpatient programs may be a wise course of action for patients and payers alike. Matching specific and timely treatment to the appropriate diagnosis makes sense for all stakeholders, despite the requisite time involved in administering the assessments.

Future directions for efficiently diagnosing mental illness in community psychiatric facilities should include assessing variables related to patients' socioeconomic factors, referral sources, age of patients, and stigma. Although currently limited in application, including biomarkers and imaging data to make clinical diagnoses will also help determine cost-effective and practical structured diagnostic tools that busy community clinicians may implement in their treatment planning for patients.

REFERENCES

1. Beck AT, Ward CH, Mendelson M, et al. Reliability of psychiatric diagnosis 2: a study of consistency of clinical judgments and ratings. *Am J Psychiatry*. 1962;119:351–357.
2. Wing JK, Birley JL, Cooper JE, et al. Reliability of a procedure for measuring and classifying "present psychiatric state." *Br J Psychiatry*. 1967;113:499–515.
3. Spitzer RL, Endicott J. *Current and Past Psychopathology Scales*. New York: New York State Psychiatric Institute, Biometrics Research Division; 1968.
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 3rd edition*. Washington, DC: American Psychiatric Press, Inc.; 1980.
5. Amador XF, Flaum M, Andreasen NC, et al. Awareness of illness in schizophrenia and schizoaffective and mood disorders. *Arch Gen Psychiatry*. 1994;51:826–836.
6. Bowie CR, Twamley EW, Anderson H, et al. Self-assessment of functional status in schizophrenia. *J Psychiatr Res*. 2007; 41:1012–1018.
7. Burdick KE, Endick CJ, Goldberg JF. Assessing cognitive deficits in bipolar disorder: are self-reports valid? *Psychiatry Res*. 2005;136:43–50.
8. Harvey PD, Heaton RK, Carpenter, WT, Jr., et al. Diagnosis of schizophrenia: consistency across information sources and stability of the condition. *Schizophr Res*. 2012;140:9–14.
9. Schwartz JE, Fennig S, Tanenberg-Karant M, et al. Congruence of diagnoses 2 years after a first-admission diagnosis of psychosis. *Arch Gen Psychiatry*. 2000;57:593–600.
10. Bromet EJ, Naz B, Fochtmann LJ, et al. Long-term diagnostic stability and outcome in recent first-episode cohort studies of schizophrenia. *Schizophr Bull*. 2005;31:639–664.
11. Yutzy SH, Woofert CR, Abbott CC, et al. The increasing frequency of mania and bipolar disorder: causes and potential negative impacts. *J Nerv Ment Dis*. 2012;200:380–387.
12. Keck PE Jr, Kessler RC, Ross R. Clinical and economic effects of unrecognized or inadequately treated bipolar disorder *J Psychiatr Pract*. 2008;14(Suppl 2):31–38.
13. Hoch P, Pollatin P. Pseudoneurotic forms of schizophrenia. *Psychiatric Q*. 1949;23:248–276.
14. Pogge DL, Wayland-Smith D, Zaccario M, et al. Diagnosis of manic episodes in adolescent inpatients: structured diagnostic procedures compared to clinical chart diagnoses. *Psychiatry Res*. 2001;101:47–54.
15. Zimmerman M, Ruggero CJ, Chelminski I, Young D. Is bipolar disorder overdiagnosed? *J Clin Psychiatry*. 2008;69:935–940.
16. Hirschfeld RM, Vornik LA. Recognition and diagnosis of bipolar disorder. *J Clin Psychiatry*. 2004;65(Suppl 15):5–9.
17. Leboyer M, Kupfer DJ. Bipolar disorder: new perspectives in health care and prevention. *J Clin Psychiatry*. 2010;71:1689–1695.
18. Harvey PD, Yehuda R. Strategies to study risk for Post-traumatic stress disorder. In Yehuda R (ed). *Risk Factors for Post-traumatic Stress Disorder*. Washington, DC: American Psychiatric Press, Inc.;1999:1–22.
19. Jäger M, Haack S, Becker T, Frasch K. Schizoaffective disorder: an ongoing challenge for psychiatric nosology. *Eur Psychiatry*. 2011;26:159–165.
20. Malhi GS, Green M, Fagiolini A, et al. Schizoaffective disorder: diagnostic issues and future recommendations. *Bipolar Disord*. 2008;10:215–230.
21. Lake CR, Hurwitz N. Schizoaffective disorder merges schizophrenia and bipolar disorders as one disease: there is no schizoaffective disorder. *Curr Opin Psychiatry*. 2007;20:365–379.
22. First MB, Spitzer RL, Gibbon M, et al. *User's guide for the Structured Clinical Interview for DSM-IV Axis I (SCID-I)*. Washington, DC: American Psychiatric Press, Inc.;1995.
23. Sheehan DV, Lecrubier Y, Harnett-Sheehan K, et al. The Mini International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview. *J Clin Psychiatry*. 1998;59(Suppl 20):22–33. ■